DOCUMENT RESUME

ED 293 860 TM 011 426

AUTHOR Redfield, Doris L.
TITLE Expected Student Achievement and the Evaluation of Teaching.
SPONS AGENCY Kentucky Career Ladder Commission.
PUB DATE Apr 88
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988). For a related document, see TM 011 425.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Academic Aspiration; Elementary Secondary Education; Evaluation Methods; Goal Orientation; *Objectives; *Student Improvement; *Teacher Evaluation; Teacher Improvement
IDENTIFIERS *Goal Setting; Kentucky Career Ladder Plan

ABSTRACT

The development of processes for considering student achievement data in the evaluation of teaching is discussed. As an alternative to the inappropriate and indefensible use of standardized test scores, the project on Expected Student Achievement (ESA) of the Kentucky Career Ladder Commission considered a management by objectives, or goal setting approach. In September 1986, 26 teachers from kindergarten through grade 12 were selected to participate in the ESA project. Goals were drafted and modified by project participants. It was proposed that participating teachers would select from four to eight goals for documentation. Near the end of the school year, each teacher met with the principal to reach agreement on the degree to which each set of Student Achievement Outcome goals had been met. Experiences of the project indicate the possibility of developing an effective and equitable system of teacher evaluation along these lines. The Goal Assessment/Documentation Forms are appended, which guided the work of the teachers, three principals, and two instructional supervisors involved in the study. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED293860

Expected Student Achievement and the Evaluation of Teaching

Doris L. Redfield

On leave from Western Kentucky University to the

U.S. Department of Education, Office of Educational Research*

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

DORIS L. REDFIELD

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Running Head: Student Achievement

Paper presented in G. Galluzzo (Symposium Organizer) Instrumentation for Evaluating Teachers for Professional Advancement: Development and Research, 1988 Annual Meetings of the American Educational Research Association, New Orleans, LA.

TM 011 426

* The work described in this paper was supported by the Kentucky Career Ladder Commission and represents the views of the author. No official support or endorsement of this work by the U.S. Department of Education is intended or should be inferred.

Expected Student Achievement and the Evaluation of Teaching

Doris L.Redfield

The focus of this paper is the development of processes for considering student achievement data in the evaluation of teaching. A limited discussion of issues and data based results is provided for contextual purposes only; details appear elsewhere (Kentucky Career Ladder Commission, 1987; Redfield, 1987; Redfield et al., 1986; Redfield & Craig, 1987a, 1987b).

Background Information

Steps 3 and 4 of Kentucky's Career Ladder Plan (Kentucky Career Career Ladder Committee, 1985) called for the evaluation of a teacher "regarding the achievement of his/her students . . . based on a determination of whether or not the students have been achieving at the expected level." However, the Kentucky Career Ladder Commission came to realize that the Kentucky Career Ladder Pilot project planned for 1986-87 could not adequately address the many complex issues surrounding the use of student achievement data in the evaluation of teaching. Hence, a special, separate project on Expected Student Achievement (ESA) was funded. The issues fueling the Commission's decision are briefly highlighted below.

Issues

Measures of student achievement are most often conceptualized as scores on standardized achievement tests. However:

- Standardized achievement tests are designed to assess students' performance, not teachers' effectiveness.

- Not all teachers teach subject matter measured by routinely administered standardized achievement tests.
- Not all teachers work with students represented by test norms.
- Expectations of student achievement may vary. For example, average performance or gain may not be a defensible expectation for non-average students (e.g., handicapped, gifted).
- When students are taught by more than one teacher, it is difficult to determine which outcomes may be uniquely attributable to any particular teacher.
- There are educational outcomes which are valued by teachers and parents but which are not typically measured using traditional standardized achievement tests.
- Not all factors influencing student achievement are under the direct control of teachers (e.g., ability, home situations).

Addressing the Issues

As an alternative to the inappropriate, indefensible use of standardized achievement test scores in the evaluation of teachers, the ESA considered a management by objectives (MBO) or goal setting approach. Using this approach, participating teachers and their principals negotiated sets of Student Achievement Outcome (SAO) goals and the degree of goal attainment. The Kentucky Career Ladder Commission recognized that the reliability and validity of a SAO goal setting approach could not be demonstrated until a system was conceptualized, developed, and tested. Hence, the ESA project implemented during the 1986-87 school year represented the first step in an ongoing, developmental process. The focus of this paper is the instrumentation and related procedures developed during, and resulting from, the work implemented during 1986-87. Proposed

plans for continuing work are summarized in the Discussion section and detailed elsewhere (Kentucky Career Ladder Commission, 1987, 1988; Redfield et al., 1986).

Developmental Processes

In September 1986, 26 teachers representing a wide variety of grade levels (K-12) and teaching areas (special education, gifted, vocational arts, visual arts, social sciences, basic skill areas, etc.) were selected for participation in the ESA project (Kentucky Career Ladder Commission, 1987; Redfield & Craig, 1987a, 1987b). The selected teacher participants, three principals, and two instructional supervisors then met for a full day with the project director. The purposes of the meeting were to: (a) introduce the group to the problems surrounding the use of student achievement data in the evaluation of teaching, (b) consider potential approaches to some of those problems, and (c) establish procedures for trying an approach to problem resolution.

The group agreed to try a SAO goal setting approach to illustrate: (a) the kinds of student outcomes they were working toward and (b) how they would evaluate the degree to which those outcomes were attained. The Goal Assessment/Documentation Forms (GADFs), shown in the Appendix, guided their work throughout the project year. The original GADFs were drafted by the project director; they were subsequently modified by project participants to reflect both their substantive and logistic concerns. It is the modified versions that are appendixed.

Goal Documentation and Assessment

In preparing to use the GADFs, project participants asked

the project director to prepare a one page synopsis of the project. They then presented this synopsis to their principals and made an appointment for a conference to discuss their participation in the project. A goal of the conference, as described in the synopsis, was for each teacher and his/her principal to negotiate a GADF for each of the teacher's proposed goals. The project synopsis also emphasized that the principal's faculty evaluation of the teacher was not to be influenced by the teacher's participation in the ESA project.

Project participants had assigned particular meaning to various terms used throughout the GADFs. Term descriptions/definitions are labeled "Explanations for Items on the Goal/Assessment Documentation Form" and appear in the Appendix.

Brief consideration of selected items from the GADFs should clarify the procedures used throughout the ESA project. Item 1 (teacher's name) served to identify each teacher's work. Items 2 and 3 were intended to provide demographic and contextual information. Project participants, as well as previously consulted teachers (Redfield, et al., 1986), were concerned with the inadequacy of traditional student achievement measures for assessing many teaching-learning situations. Hence, the variety of student types, subject matter areas, and group sizes represented by even this relatively small group of 26 teachers was documented.

Item 4 represents the participating teachers' determination that they desired at least four kinds of outcomes for their students. Some desired outcomes were described as

"academic" in nature (e.g., basic skill attainment) and some were not (e.g., positive attitudes toward learning, prosocial behavior). Whether academic or "nonacademic," some desired outcomes were considered "specific" to a particular teaching-learning situation (e.g., development of self-help skills in handicapped students) whereas others were considered "general," applying to all types of students regardless of class content or grade level (e.g., positive self-concept). Each of these four types of desired outcomes (i.e., general academic, specific academic, general nonacademic, specific nonacademic) are further explained in the Appendix.

The teachers agreed that they would each document from four to eight SAO goals, at least one from each of the four categories described above. Any goal might be short-range, mid-range, or long-range in scope. Short-range goals were defined as interum goals to be accomplished in less than the total period of time spent by a teacher with a student, group, or class (e.g., a goal targeted for accomplishment by the end of the first quarter of a semester-long class). Mid-range goals were defined as those slated for accomplishment by the end of the time period spent by a teacher with a student, group, or class. Long-range goals were defined as those worked toward, but not necessarily accomplished, during a teacher's assignment to work with a particular student, group, or class (e.g., responsibility, writing).

The goals selected for documentation by each teacher were <u>not</u> to be conjured up as a result of participating in the ESA project; rather, goals were to be selected from the repertoire of goals that each teacher had already developed or planned to

pursue throughout the school year. The importance of not changing what they would ordinarily do was emphasized because an objective of the project was to document what teachers reasonably do to demonstrate their students' achievements, especially when standardized test scores cannot be appropriately used.

Item 5 called for a statement of the teacher's goal. The greatest difficulty teachers seemed to encounter was stating their goals in operational terms. In such cases, project staff provided technical assistance.

The intent of Item 6 was to document the variety of sources teachers draw upon in determining what students need to know. An extreme finding was that of the 111 goals documented throughout the project year, 74 (67%) had a basis in some sort of "professional judgment" on the part of the teacher; only 2 of the 111 goals were based on consultation with other professionals or colleagues.

Item 7 was included to address the concern that some teachers might identify trivial goals for any number of reasons, (e.g., easily attainable; dictated by a particular, arbitrarily selected curriculum). Basically, teachers wanted the significance of their work considered in the evaluation process. The goals documented throughout the ESA project yielded a mean value (across all goals and all teachers) of 4.50 on a 5.00 scale, with 1 being insignificant and 5 being highly significant.

Item 8 was included so that the evaluation process might take the difficulty of reaching any particular goal into account. The teacher participants emphasized that unless goal

difficulty was considered, teachers might avoid selecting important goals simply because they could be difficult to fully attain and, hence, result in unfairly low evaluations.

The purpose of Item 9 was to gather information regarding factors hypothesized as influencing the difficulty level of each goal. This information might be used: (a) in the determination of appropriate covariates if ultimate scoring procedures are based on regression modeling and (b) in future efforts to develop and calibrate a bank of goals from which teachers might select designated quantities and/or types of goals.

Item 10 required teachers to designate the type(s) of documentation they would gather to demonstrate progress toward each of their goals. Here, teachers were quite creative. In fact, sometimes talking them through item 10 helped them operationalize their goal statements. A pertinent finding was that standardized test scores, of any kind, were proposed as documentation for only 11 of the 111 documented goals. Other proposed types of documentation included charts, checklists, performance ratings, student evaluations, observation data, official records (e.g., attendance), task completion, and grades.

Item 11 was included to encourage consideration of the validity of the proposed forms of documentation. This item proved difficult for the teachers and their principals and was seemingly related to their difficulty in operationalizing the goal statements. The rating assigned to this item called for a rationale that might be used in future development efforts. For example, a menu of valid procedures for assessing particular

goals (perhaps selected from a menu) might be developed.

The intent of item 12 was to document the times during which teachers collected data for showing progress toward each of their SAO goals. It was hoped that responses would contribute to an understanding of the time and effort required by various documentation procedures and might also have implications for the training needs of teachers, principals, et al. As might be expected, the nature of particular goals often determined the optimal or most efficient time for collecting evidence of progress or goal attainment. For example, mid-range academic goals might be efficiently monitored via pretesting at the beginning of a semester or year and posttesting at the end of a semester or year. However, monitoring progress toward specific objectives necessary for meeting a mid-range goal might require monitoring at the end of each instructional unit. Teachers varied greatly in their specification of times for collecting documentation. Examples of the data collection schedules adopted by the teachers included: as necessary; beginning and/or throughout and/or ending of a week, month, unit, semester, etc.; each class, day, week, month, etc.; and/or after a specific event (e.g., after a test).

If documentation of goal progress is to be assessed, the data must take on an interpretable form. For example, it is difficult to defensibly interpret the meaning of a notebook containing a student's writing assignments. It is relatively easy to defensibly interpret the meaning of a list of scores representing a student's performance on each of those same writing assingments when criteria for scoring are clearly

specified. To encourage the assignment of meaning to their collected data, the teacher participants restricted themselves to providing but one page of documentation per goal. A second reason for this restriction was to cut down on paperwork. However, it was soon discovered that less paper did not mean less work (or time)! Item 13 asked teachers to specify how they assigned meaning to the data collected for documentation purposes.

Item 14 was included to document procedures used by practicing teachers to enhance the fairness (i.e., lack of positive or negative bias) of their assessments. Teachers' responses to this item included: allowing adequate time for students to learn material and prepare for exams, protecting student anonymity, averaging several scores obtained at various times rather than depending on one score to represent overall achievement, providing clear instructions and expressions of expectations, predetermining and announcing grading criteria, and using assessment techniques deemed valid (by the teachers and their principals) for the purpose at hand.

A task of the ESA project involved consideration of what constitutes fair expectations of student achievement. Reasonable expectations might well be expected to differ across student types (e.g., handicapped vs. gifted), teaching-learning domains (e.g., basic skills vs. behaviors vs. attitudes and affects), and grade levels. Therefore, item 15 attempted to document what constituted expected student achievement for the teaching-learning context represented by each goal. As anticipated, levels of expectation differed from teacher to

teacher and from goal to goal according to any given situation. The criteria for expected achievement were stated by teachers in terms of: designated amounts of change in performance from one point in time to another, competitive acceptance rates (e.g., in art shows), levels of conformity or compliance, grades of various kinds (e.g., points, proportions, letter grades), infractions, mastery, participation, and number or proportion of students passing any given assignment, task, or other "hurdle."

In late April or early May 1987, each participating teacher met with his/her principal to reach agreement on the degree to which each goal had been met. A GADF developed for Conference II (see Appendix) was used to guide the conference. A five-point scale, ranging from 5 (representing significant progress) to 1 (representing no progress), was used to assign the ratings. The mean ratings for individual teachers across goals ranged from 2.0 to 5.0. The grand mean across all teachers for all goals was 3.56. Teachers and principals were asked to provide a rationale for each rating. For 59 of the 111 goals (53%), the rationale was stated in terms of the relationship between the documented outcome and the criteria designated in item #15 GADF - Conference I) for expected achievement. It seems noteworthy that 18 of the 26 teacher participants provided anecdotal accounts of the outcomes associated with their numerical ratings -- as if the numbers could not tell the whole story.

Conceptually, the rating assigned to item #2 on the GADF for Conference II might be added to the corresponding ratings assigned to items 7 and 8 on the GADF for Conference I. Then, totals might be averaged across a teacher's goals.

Discussion/Conclusions

Experiences throughout the ESA project, as described in this paper and elsewhere (e.g., Kentucky Career Ladder Commission, 1988; Redfield, 1987; Redfield & Craig, 1987a) have numerous implications for continuing development of a multipurpose teacher evaluation system that includes consideration of student outcomes. The argument concerning student achievement and teacher evaluation is not whether student achievement should be included; rather, the issue is fair and defensible inclusion. Additionally, there seems to be increasing awareness that a viable system must meet both summative and formative evaluation needs. Such a multipurpose evaluation system will require careful attention to the training and support needs of both evaluators and evaluatees. For example, teachers and their evaluators would at least require: (a) training in measurement and conferencing/negotiation skills and (b) ready access to technical expertise. Teacher Education programs would, in many cases, require re-focusing to help meet the needs of practicing educators as well as the needs of teacher candidates (see Redfield, 1988).

In order to continue the development and testing of a multipurpose teacher evaluation system that both calls for teacher accountability and allows for professional development, at least the following events would need to occur over a continuing two to four year period of time.

- Determine if a relatively large number of teachers and principals, given adequate training and support, are able to negotiate SAO goals and appropriate assessments for goal attainment.
- Determine if this relatively large number of

teachers and principals could provide a sufficient variety of SAO goals and assessment techniques for the development of a menu from which core goals and assessment techniques could be validated against professional concensus.

- Determine the role of "specific" (vs. "general") goals as defined by the ESA project, in the evaluation system.
- Determine the number of teachers with whom principals or other supervisors/evaluators could reasonably work.
- Test a system for taking SAO goal significance and difficulty into account.
- Determine the degree to which the process is able to differentiate good teachers from the best teachers.
- Develop and test an appeals process.
- Determine how to provide school personnel with the ongoing support needed to maintain development efforts to enhance SAOs.
- Develop and test instruments for specifying, documenting, and evaluating SAO goals.
- Develop and test training programs for teachers and the supervisors responsible for assisting and/or evaluating them.

References

Kentucky Career Ladder Commission. (1988). Research report: Kentucky Career Ladder Pilot Study. Available from J.R. Craig, College of Education and Behavioral Sciences, Western Kentucky University, Bowling Green, KY 42101.

Kentucky Career Ladder Commission. (1987). Technical report: Special project on expected student achievement. Available from D.L. Redfield, U.S. Department of Education, Office of Research, 555 New Jersey Ave. N.W./Room 608, Washington, DC 20208.

Kentucky Career Ladder Committee. (1985). Report of Kentucky Career Ladder Committee: Kentucky Career Ladder Plan for Teachers ("Bluebook"). Available from R.S. Pankratz, College of Education and Behavioral Sciences, Western Kentucky University, Bowling Green, KY 42101.

Redfield, D.L. (1987, November). Expected student achievement as a potential factor for assessing teacher effectiveness. In R.S. Pankratz (Symposium Organizer), The Assessment of Teacher Performance Using Classroom Observations, Structured Interviews, Goal Attainment and Expected Achievement: Lessons Learned, Annual Meetings of the Mid-South Educational Reseaarch Association, Mobile, AL.

Redfield, D.L. (1988). Pupil achievement: The weakest link in the evaluation chain. In W.J. Gephart & J.B. Ayers, Teacher Education Evaluation, Kluwer Academic Publishers, Boston, MA.

Redfield, D.L. et al. (1986). Report prepared for the Kentucky Career Ladder Commission: Use of expected student achievement data in teacher evaluation. Available from D.L. Redfield, U.S. Department of Education, Office of Research, 555 New Jersey Ave. N.W./Room 608, Washington DC 20208.

Redfield, D.L., & Craig, J.R. (1987a, October). The use of pupil outcome measures in the assessment of teaching when reliance upon standardized test scores is not defensible. In D.L. Redfield (Panel Organizer), Perspectives on the Use of Student Achievement Data in the Evaluation of Teaching, Annual Meetings of the American Evaluation Association, Boston, MA.

Redfield, D.L., & Craig, J.R. (1987b, November). Identifying and documenting student outcomes for use in the evaluation of teachers when standardized achievement tests do not apply. In D.L. Redfield (Symposium Organizer), The Hows and Whys of Using a Variety of Student Achievement Data to Assess the Effects of Teaching and Schooling: The State of the Mid-South Region, Annual Meetings of the Mid-South Educational Research Association.

APPENDIX

Goal #: _____

GOAL/ASSESSMENT DOCUMENTATION FORM (GADF): Conference I

1. Teacher:

2. Target class(es)/group(s) -- specify grade, student type, content (e.g, 9th-grade, required, Civics; 4th-grade, self-contained; high school, elective, Art)

3. Number of targeted students:

4. Type of goal (check all that apply):

_____ specific	_____ academic	_____ short-range
_____ general	_____ nonacademic	_____ mid-range
		_____ long-range

5. Goal statement:

6. Source of goal (check all that apply):

_____ essential skills list	_____ (textbook) scope & sequence
_____ state curriculum guide	_____ professional literature
_____ coursework	_____ personal belief
_____ professional association guidelines	
_____ other (specify):	

7. Educational significance of the goal (circle one number):

1 2 3 4 5

insignificant highly significant

Because:

8. Ease of goal attainment (circle one number):

1 2 3 4 5

very easy very difficult

Because:

(OVER)

9. Factors influencing the ease of goal attainment

_____ SES (describe):

_____ ability (describe):

_____ other (specify/describe):

10. What information will be gathered to document the degree to which the goal is achieved?

Relationship between the goal and the proposed documentation (circle one number):

1 2 3 4 5

poor superior

Because:

12. When will the documenting information be gathered?

13. How will weights (values, labels) be assigned to the documentating information?

14. What steps will be taken to enhance the fairness and defensibility of the information gathered and the weights assigned to it?

15. The weights assigned to the gathered information will be interpreted as follows.

o no progress toward the goal:

o less than expected progress:

o expected progress:

o progress slightly above expectation:

o progress significantly above expectation:

16. Date of Conference I:

17.

Points of discussion/ disagreement	Nature of discussion/ disagreement	Outcome of discussion/ disagreement

Notes

Goal #: _____

GOAL/ASSESSMENT DOCUMENTATION FORM (GADF): Conference II

1. Teacher:

2. Based upon the documenting information gathered, the weights (labels/values) assigned to it, and the interpretation of those weights, progress toward the goal may be best described as (circle one):

1 2 3 4 5

no progress significant progress

Because:

3. Date of Conference II:

4.

Points of discussion/ disagreement	Nature of discussion/ disagreement	Outcome of discussion/ disagreement

Notes

EXPLANATIONS FOR ITEMS ON THE GOAL/ASSESSMENT DOCUMENTATION FORM

Target Class(es)/Group(s)

The target class or group is the group of students toward whom the goal is directed. This group may be an entire class, a subgroup of a class, an individual student, or several classes combined.

Type of Goal

SPECIFIC goals are goals that are unique to one particular teacher. They are goals that are not likely to apply to other teachers at the same grade level or within the same content area. GENERAL goals are goals that are likely to be goals of all teachers regardless of grade level or content area. ACADEMIC goals are aimed at increases in cognitive knowledge, academic achievement, or skill development. NONACADEMIC goals are not related to academic content and generally concern affective or behavioral outcomes. SHORT-RANGE goals are interum goals to be accomplished during a period of time less than the total period of time a given teacher spends with a given class/group (e.g., a semester goal when the teacher has students for a year; a unit goal when a teacher has students for a quarter). MID-RANGE goals are goals to be accomplished by the end of the total period of time a given teacher spends with a given class/group (e.g., end-of-year, end-of-semester for semester length classes). LONG-RANGE goals are those which are worked toward, but which may not be fully accomplished, within the period of time a given teacher works with a given group of students (e.g., independent learning).

Goal Statement

The goal statement is simply a statement of a goal or an objective that will be worked toward. Stating each goal as a performance objective (i.e., by describing the circumstances under which the goal

will be accomplished, what students will be expected to do, and the criterion for successful performance) should facilitate its clear understanding and accomplishment.

Documentation (see item # 10)

Documentation refers to the data that will be collected as evidence that a particular goal has or has not been achieved. If goals are stated as performance objectives, then documentation refers to the measure of performance used by the teacher. Documentation may include test/quiz scores, grades, observation checklists, annecdotal records, videotapes, etc.

Relationship Between Goal and Proposed Documentation (item # 11)

The issue, here, is the degree to which the collected documentation is appropriate for assessing goal attainment (i.e., fairness, reliability, and validity of measures used by teachers in assessing students). For example, a teacher-made test regarding nutrition facts may provide an appropriate measure of nutrition facts but not of physical fitness. Therefore, if the goal is to train physical fitness, then the match between goal and documentation (nutrition test) is "inappropriate;" however, if the goal is that students learn facts about nutrition, then the goal-documentation match is "satisfactory." Assessing the degree to which students apply nutrition facts in menu planning might be considered a "better than normally expected," goal-documentation match, and observing students' lunch selections one day per week for five weeks might be considered an "extremely appropriate" documentation-goal match.

When Documentation will be Gathered (item # 12)

Here, the points at which documentation will be gathered should be designated (e.g., every Friday, at the end of each unit, once at the

end of the school year).

Weight(s) (item #13)

Weight refers to the description the teacher assigns to each piece of documentation. The weight might be a letter grade, percent correct, number correct, plus vs. check vs. minus mark, smiley vs. frowny face, standardized test score, etc.

Fairness/Defensibility (item #14)

The steps taken by the teacher to enhance the fairness of the gathered documentation are noted here. Examples might include using tests with established reliability, blindly scoring papers, using standard evaluation procedures, etc.

Interpretation of Weights Assigned to Documentation (item # 15)

This item requires describing what the weights assigned to the documentation mean. For example, a weight of "A" might indicate that progress toward the goal was significantly exceeded. On the other hand, if the goal is mastery learning, then a weight of 90% correct might reflect expected progress.

Points/Nature/Outcome of Discussions/Disagreements (item # 17)

Note those items that generated discussion or disagreement, why there was discussion or disagreement, and the end result of each discussion or disagreement.